Assessment of Visual Association Memory in Low-Educated, Non-Western Immigrants with the Modified Visual Association Test

Sanne Franzen a, Esther van den Berg a, Yuled Kalkisim a, Lotte van de Wiel b, Marleen Harkes c, Rozemarijn L. van Bruchem-Visser d, Frank Jan de Jong a, Lize C. Jiskoot a, Janne M. Papma a

aDepartment of Neurology, Erasmus MC University Medical Center, Rotterdam, The Netherlands; bDepartment of Medical Psychology, Maasstad Ziekenhuis, Rotterdam, The Netherlands; cDepartment of Geriatric Medicine, Maasstad Ziekenhuis, Rotterdam, The Netherlands; dDepartment of Internal Medicine, Erasmus MC University Medical Center, Rotterdam, The Netherlands

Keywords
Alzheimer’s disease · Dementia · Neurodegenerative diseases · Cross-cultural comparison · Literacy · Education · Cognition

Abstract
Background: Neuropsychological tests are influenced by culture, language, level of education, and literacy, but there are few cognitive tests of which the applicability in ethnic minority populations has been studied. Objectives: The aim of this study was to assess the reliability and validity of the Visual Association Test (VAT), a test of visual association memory, in a non-Western, low-educated memory clinic population. Additionally, a modified version of the VAT using colored photographs instead of line drawings was studied (mVAT). Method: Both the original VAT and the mVAT were administered to non-Western immigrants (n = 73) from 2 multicultural memory clinics in Rotterdam, The Netherlands, and a control sample of non-demented Turkish elderly (n = 14) with low education levels (32 and 29% illiterate, respectively). Results: Both the VAT and the mVAT were able to discriminate persons with and without dementia (area under the curve: VAT, 0.77–0.88; mVAT, 0.85–0.95). The mVAT had more homogeneous item difficulty levels than the VAT. Administration of parallel versions of the VAT and the mVAT within the same person revealed higher scores on the mVAT (Z = –3.35, p = 0.001). Conclusions: The mVAT is a reliable and valid measure of memory in non-Western immigrants. Clinicians and researchers should be aware that the memory performance of immigrants may be systematically underestimated when using tests with black-and-white line drawings, such as the original VAT.
Introduction

Neuropsychological examination is fundamental to the assessment of dementia. Neuropsychological test performance is known to be substantially affected by culture, language, (quality of) education, and literacy [1–8]. For example, healthy illiterate persons have lower scores on visual naming tests using black-and-white line drawings than literate persons [9–12]. This difference disappears when colored photographs are used [10], which is most likely related to the higher level of detail provided by the colored photographs [13]. This example illustrates how tests developed for educated, Western people cannot readily be used in other populations. In recent years, new screening instruments for dementia, such as the Cross-Cultural Dementia Screening or CCD [14] and the Rowland Universal Dementia Assessment Scale or RUDAS [15–18], have therefore been developed and validated for ethnic minority populations. These instruments are designed to screen for dementia. However, domain-specific neuropsychological tests that can determine the underlying etiology are lacking.

The Visual Association Test (VAT) [19] is a test of visual association memory that is particularly able to discriminate between Alzheimer’s disease (AD) and other types of dementia, as performance is associated with atrophy of the medial temporal lobe [20]. In the VAT, patients are required to remember two interacting objects (such as a stroller with a bird in it, or a monkey with an umbrella) presented in the form of black-and-white line drawings. Although this test is frequently used in clinical practice and was previously recommended as best practice in the neuropsychological assessment of non-Western immigrants [21], the reliability and validity of the VAT have not been assessed in an ethnic minority population or in people who are illiterate.

Given the above-mentioned difficulties regarding the naming of black-and-white line drawings in healthy illiterate people, the question can be raised of whether the VAT puts illiterate or low-educated patients at a disadvantage, for example due to difficulties recognizing and thus remembering the objects. These difficulties may also pose a threat to the validity of the VAT as a measure of visual association memory in these persons. We hypothesize that people who are low educated could benefit from a test that uses colored photographs instead of black-and-white line drawings. The aims of this study were to examine the reliability and validity of the original version of the VAT and a modified version using colored photographs (mVAT) and to compare performance on both tests in a population of low-educated non-Western immigrants.

Materials and Methods

Participants

Seventy-three non-Western immigrant patients who visited the outpatient multicultural memory clinics of the Erasmus MC University Medical Center and the former Havenziekenhuis in Rotterdam, The Netherlands, were included between April 2016 and October 2018. The patients had immigrated from Turkey (n = 34), Morocco (n = 13), Cape Verde (n = 10), Pakistan (n = 4), Iraq (n = 2), Afghanistan (n = 2), the State of Palestine (n = 1), Syria (n = 1), Egypt (n = 1), China (n = 1), Venezuela (n = 1), the Dutch Antilles (n = 1), Suriname (n = 1), and Macedonia (n = 1). The diagnostic workup consisted of a comprehensive clinical evaluation, with history taking by a geriatrician or neurologist, a neuropsychological assessment, laboratory screening with blood tests, and (in a subset of patients) structural brain imaging (computed tomography, n = 23; magnetic resonance imaging, n = 25). Diagnosis was determined in a multidisciplinary consensus meeting which included a neuropsychologist, a neurologist, a radiologist, and a geriatrician, using the diagnostic research criteria for dementia subtypes [22–24].

Fourteen healthy Turkish community-dwelling individuals were included as a control group. The inclusion criteria for the control group were: age ≥50 years, free of self-reported cognitive complaints, and a RUDAS [15] score ≥23. For 9 healthy controls, informants filled out the short IQ-code [25], confirming the
patient’s self-reported absence of cognitive complaints. In the other cases, no informant was available. All healthy controls provided a written informed consent. They were recruited in an urban area through centers providing activities for Turkish elderly and through the personal networks of included participants.

**Measures**

**The Original VAT**

In the original VAT, patients are asked to name 6 stimuli on consecutive black-and-white line drawings (the cue cards) presented in a paper booklet. Patients are then shown cards on which the previous stimuli are interacting with new stimuli, and they are asked to name both items on each consecutive card. Patients are then again shown the initial cue cards and asked to name the missing item. This procedure is repeated in trial 2 (unless there is a maximum score on the first trial). Naming errors are allowed as long as the names are specific enough to be identified as correct or incorrect in the reproduction trial, i.e., “prickly animal,” “prickly thing,” or even “brush” would be sufficiently specific to indicate a hedgehog and would thus be considered correct.

**Modification of the VAT**

To create a photograph version of the VAT, henceforth called the modified VAT (i.e., mVAT), all black-and-white line drawings from the original VAT were replaced with colored photographs of the same objects (Fig. 1). No other changes were made to the test. Stimuli were presented on A5-sized booklets, similar to the original VAT. The original VAT has multiple parallel versions. Therefore, modifications of both the original VAT version A and the parallel version B were made.

**Other Cognitive Tests**

The patients were tested with the CCD [14], a comprehensive screening test for dementia, assessing the domains of memory, mental speed, and executive functioning. It was specifically developed and validated for a large sample of immigrants in The Netherlands, and normative data are available. Furthermore, patients were administered the RUDAS [15] and/or the Mini Mental State Examination (MMSE) [26]. In some patients, other tests, such as the Location Learning Test [27] (n = 15) and category verbal fluency (supermarket) (n = 10), were administered as well.

**Procedure**

All participants from the outpatient clinic underwent a neuropsychological examination. Interpreters were present during these examinations, as well as during the intake interview, for all participants who were insufficiently fluent in Dutch (96%, n = 70). The administration procedures for the VAT (and thus the mVAT) according to the test manual [19] were followed, including prorating of the second trial. The administration time of the VAT (and the mVAT) is approximately 5 min. As the stimuli were identical for both the VAT and the mVAT (apart from line drawings vs. photographs), patients could not be administered version A of both the VAT and the mVAT. Therefore, all patients were administered either VAT version A and mVAT version B
or VAT version B and mVAT version A. In the majority of cases, the mVAT was administered before the VAT to ensure that a higher score on the mVAT was not caused by the participant being aware at the first trial that the stimuli needed to be remembered. In a subset of patients, this order was reversed.

Control participants were assessed in Turkish, either at home or at the day activity center, by a neuropsychologist who is a native speaker of Turkish (Y.K.). Similar to the patients, the order of administration of the VAT versus the mVAT, as well as the versions that were used, was varied.

### Statistical Analysis

Differences in demographic characteristics were analyzed with \( \chi^2 \) tests for nominal data and \( t \) tests for continuous data. Cronbach’s \( \alpha \) and item-total correlations were used to determine the internal consistency, and Cochran \( Q \) tests were used to assess item difficulty levels. For reliability analyses, both parallel versions of the VAT and the mVAT were analyzed separately. For all other analyses, parallel versions A and B of the original and A and B of the mVAT were merged. As the scores of both the VAT and the mVAT were not normally distributed, a Wilcoxon signed-rank test was used to analyze intraindividual differences in the performance on the mVAT versus the VAT. Spearman correlations were used for correlational analyses to determine the convergent validity and the relationship with demographic variables. Receiver operating characteristic curves were used to assess diagnostic accuracy. In secondary analyses differences were analyzed in participants with a low education level (less than primary school) versus participants who had received more education (primary school and up).

### Results

As is shown in Table 1, controls were more often female, but the groups did not differ in terms of age, years of residence in The Netherlands, or level of education. For 14% of the patients in the sample, the diagnosis could not be determined definitively with the current

### Table 1. Demographic characteristics, cognitive test scores, and group comparisons for the whole sample

<table>
<thead>
<tr>
<th></th>
<th>Controls (n = 14)</th>
<th>Memory clinic patients (n = 73)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>62.21±11.49</td>
<td>68.48±11.00</td>
<td>ns</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 years of education/illiterate</td>
<td>4 (29)</td>
<td>23 (32)</td>
<td>ns</td>
</tr>
<tr>
<td>1 year of education up to primary education</td>
<td>6 (43)</td>
<td>27 (37)</td>
<td></td>
</tr>
<tr>
<td>&gt;primary education</td>
<td>4 (29)</td>
<td>21 (29)</td>
<td></td>
</tr>
<tr>
<td>Male gender, %</td>
<td>21</td>
<td>55</td>
<td>( p = 0.02 )</td>
</tr>
<tr>
<td>Time in the Netherlands, years</td>
<td>36.9±16.3</td>
<td>38.5±7.4</td>
<td>ns</td>
</tr>
<tr>
<td>RUDAS</td>
<td>26.86±1.92</td>
<td>22.1±5.0(^a)</td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td>–</td>
<td>16.6±5.8(^b)</td>
<td></td>
</tr>
<tr>
<td>CCD objects test A(^c)</td>
<td>–</td>
<td>113.2±9.6</td>
<td></td>
</tr>
<tr>
<td>CCD objects test B(^c)</td>
<td>–</td>
<td>104.7±12.2</td>
<td></td>
</tr>
<tr>
<td>Diagnosis, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjective memory complaints</td>
<td>–</td>
<td>13 (18)</td>
<td></td>
</tr>
<tr>
<td>Mild cognitive impairment</td>
<td>–</td>
<td>13 (18)</td>
<td></td>
</tr>
<tr>
<td>Dementia (AD, VaD, mixed, and other)</td>
<td>–</td>
<td>22 (30)</td>
<td></td>
</tr>
<tr>
<td>Primary psychiatric disorder (e.g., depression)</td>
<td>–</td>
<td>8 (11)</td>
<td></td>
</tr>
<tr>
<td>Cognitive disorder due to another known medical condition</td>
<td>–</td>
<td>7 (10)</td>
<td></td>
</tr>
<tr>
<td>Could not be determined</td>
<td>–</td>
<td>10 (14)</td>
<td></td>
</tr>
</tbody>
</table>

\( \alpha \): vascular dementia. Values are displayed as means ± SD or numbers (%) unless otherwise specified. \(^a\) n = 26 patients. \(^b\) n = 54 patients. \(^c\) The maximum score for the objects test A (immediate recognition) and B (delayed recognition) is 122. The general cut-offs for dementia are ≤118 for objects test A and ≤109 for objects test B.
diagnostic procedures, or additional procedures to determine the diagnosis, such as a lumbar puncture, failed or were refused by the patient. These patients remained in the analyses. All of the included patients and controls had normal or corrected-to-normal vision. None of the patients or controls reported color blindness.

Reliability

Internal Consistency

The first trial of VAT version A displayed a Cronbach’s α of 0.75, while the first trial of VAT version B had a Cronbach’s α of 0.41 (Table 2). Cronbach’s α for the first trial of mVAT version A was 0.63, and Cronbach’s α was 0.81 for the first trial of mVAT version B. Item-total correlations were strong for the mVAT, as well as for VAT version A (Table 2). VAT version B, however, had 2 items with nonsignificant correlations with the total score for trial 1.

Item Difficulty

Table 3 shows the percentage of people who remembered an item correctly at the first trial of the VAT and mVAT versions A and B. Overall, the item difficulty was lower for the mVAT than for the VAT, without reaching ceiling effects. Item difficulty levels were homogeneous for VAT version A and mVAT version A. Heterogeneity in item difficulty levels was observed for VAT version B (Cochran’s Q = 24.8, d.f. = 5, p < 0.001). Version B of the mVAT did not have homogeneous difficulty levels either (Cochran’s Q = 22.0, d.f. = 5, p = 0.001), as item 5 (leaf-syringe) was more difficult than the other items, but after deleting this item the test was homogeneous (Cochran’s Q = 0.7, p = ns).

**Table 2.** Item-total Spearman’s correlations and Cronbach’s α for the first trial of all versions

<table>
<thead>
<tr>
<th></th>
<th>Trial 1 mVAT A (n = 32)</th>
<th>Trial 1 mVAT B (n = 55)</th>
<th>Trial 1 VAT A (n = 55)</th>
<th>Trial 1 VAT B (n = 25)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1 (ρ)</td>
<td>0.64</td>
<td>0.77</td>
<td>0.61</td>
<td>0.57</td>
</tr>
<tr>
<td>Item 2 (ρ)</td>
<td>0.57</td>
<td>0.59</td>
<td>0.59</td>
<td>0.60</td>
</tr>
<tr>
<td>Item 3 (ρ)</td>
<td>0.54</td>
<td>0.74</td>
<td>0.72</td>
<td>0.50</td>
</tr>
<tr>
<td>Item 4 (ρ)</td>
<td>0.46</td>
<td>0.73</td>
<td>0.68</td>
<td>0.34 (ns)</td>
</tr>
<tr>
<td>Item 5 (ρ)</td>
<td>0.56</td>
<td>0.73</td>
<td>0.68</td>
<td>0.69</td>
</tr>
<tr>
<td>Item 6 (ρ)</td>
<td>0.67</td>
<td>0.68</td>
<td>0.71</td>
<td>0.18 (ns)</td>
</tr>
<tr>
<td>Cronbach’s α</td>
<td>0.63</td>
<td>0.81</td>
<td>0.75</td>
<td>0.41</td>
</tr>
</tbody>
</table>

All correlations are significant at p < 0.05 unless otherwise specified.

**Table 3.** Percent correct per item of mVAT A and B and VAT A and B

<table>
<thead>
<tr>
<th></th>
<th>mVAT A trial 1</th>
<th>VAT A trial 1</th>
<th>mVAT B trial 1</th>
<th>VAT B trial 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A item 1</td>
<td>69</td>
<td>55</td>
<td>62</td>
<td>64</td>
</tr>
<tr>
<td>A item 2</td>
<td>75</td>
<td>54</td>
<td>68</td>
<td>28</td>
</tr>
<tr>
<td>A item 3</td>
<td>72</td>
<td>64</td>
<td>66</td>
<td>44</td>
</tr>
<tr>
<td>A item 4</td>
<td>60</td>
<td>44</td>
<td>64</td>
<td>72</td>
</tr>
<tr>
<td>A item 5</td>
<td>75</td>
<td>56</td>
<td>36</td>
<td>64</td>
</tr>
<tr>
<td>A item 6</td>
<td>53</td>
<td>56</td>
<td>64</td>
<td>24</td>
</tr>
</tbody>
</table>

Values are presented as percents.
For the distribution of the scores in the first trial for patients and controls, see Figure 2. There were 324 correct responses in the first trial of the mVAT \((n = 87)\); of the subjects who also completed trial 2, there were 304 cases followed by a correct response on the same item in the second trial (93.8%). Similarly, on the VAT, 260 correct responses were recorded for the first trial \((n = 81)\), 250 cases (96.2%) of which were followed by a correct response on the same item in the next trial.
Validity
  Intraindividual Performance on the VAT and the mVAT
  The intraindividual performance of all of the participants in the first trial of the mVAT was significantly higher than in the first trial of the VAT ($Z = -3.35$, $p = 0.001$). The results were comparable when the controls were analyzed separately ($Z = -2.31; p = 0.021$) or the patients separately ($Z = -2.83; p = 0.005$).

  Discriminative Validity
  The discriminative abilities of the mVAT and VAT (first trial) for patients with dementia versus healthy controls are displayed in Figure 3. The area under the curve (AUC) was 0.95 for the mVAT ($n = 36$) and 0.88 for the VAT ($n = 30$). When discriminating patients with dementia from the entire sample (including healthy controls), the AUC for the first trial of the mVAT was good (i.e., 0.85, $n = 87$), and the AUC for the first trial of the VAT was fair (i.e., 0.77, $n = 80$; Figure 3). A separate analysis of only the Turkish participants revealed similar results. The sample sizes for this study were too small to examine specific dementia subtypes.

  Convergent Validity
  The mVAT was moderately correlated with the MMSE ($\rho = 0.51$, $p < 0.001$, $n = 54$), the RUDAS ($\rho = 0.61$, $p < 0.001$, $n = 39$), and the CCD objects test part A ($\rho = 0.58$, $p < 0.001$, $n = 71$) and part B ($\rho = 0.63$, $p < 0.001$, $n = 70$). The VAT showed similarly moderate correlations with these cognitive screening measures (MMSE: $\rho = 0.54$, $p < 0.001$, $n = 49$; RUDAS: $\rho = 0.57$, $p < 0.001$, $n = 38$; CCD objects test part A: $\rho = 0.57$, $p < 0.001$, $n = 65$; and CCD objects test part B: $\rho = 0.54$, $p < 0.001$, $n = 64$). After splitting the group into a low-educated group and a highly educated group, the correlations of the mVAT with the CCD and MMSE remained similar, but the correlation with RUDAS was no longer significant in the low-educated group (Table 4).

  Relationship with Demographic Variables
  The mVAT was weakly correlated with education level ($\rho = 0.25$, $p = 0.02$) and uncorrelated with gender ($\rho = 0.06$, $p = \text{ns}$), as was the VAT (education: $\rho = 0.29$, $p = 0.01$; gender: $\rho = 0.01$, $p = \text{ns}$). The VAT was moderately correlated with age ($\rho = -0.43$, $p < 0.001$) and weakly correlated with years of residence in The Netherlands ($\rho = -0.38$, $p < 0.001$), and the mVAT was weakly correlated with age ($\rho = -0.37$, $p < 0.001$) and years of residence in The Netherlands ($\rho = -0.32$, $p < 0.01$). After splitting the group into a low-educated and higher educated group, the correlation of the mVAT with education level and with years of residence in The Netherlands only remained significant for the low-educated group (Table 4). The correlations with age and gender remained similar.
Discussion

This study showed that both the VAT and the mVAT are useful measures of visual association memory in non-Western immigrants. The mVAT, however, had a better discriminative ability and less heterogeneous item difficulty levels than the VAT.

Administration of parallel versions of the VAT and the mVAT within the same person revealed that, both for controls and for patients, higher scores were obtained on the mVAT. This finding indicates that the memory performance of non-Western immigrants may be systematically underestimated when using the original VAT with its black-and-white line drawings. This is an important finding, as memory tests for ethnic minority populations often contain visual stimuli to circumvent language barriers, and black-and-white line drawings, such as in the picture version of the Free and Cued Selective Recall Test [28] and the Location Learning Test [27, 29], are still widely used. Even some newly developed tests for illiterate and low-educated subjects contain black-and-white line drawings [30, 31].

The better performance on the mVAT is best explained by the added information about the color provided by the colored photographs (as opposed to surface texture information) [13, 32]. The added color may be particularly important for low-educated people, as the ability to decode black-and-white line drawings (such as those in the VAT) is thought to be acquired through education and literacy training [13, 33].

The reliability analyses indicated that only 1 item of the mVAT would have to be replaced before the mVAT could be applied in ethnic minority populations. This may seem counterintuitive, as some of the items of the mVAT do not necessarily appear to be universal. This finding may best be explained by the fact that misnamed or misidentified objects are not scored as incorrect as long as they are specific enough to be scored in the recall trial. Therefore, a hedgehog that is not recognized as such but is instead called a “brush” may have the same item difficulty as an item that seems more cross-culturally recognizable.

A remarkable finding was that performance on the mVAT was associated with education in a group of illiterate to minimally educated persons but not in a group with primary education levels and higher. This is in line with findings that even 1 year of education may substantially alter the performance on cognitive tests [7] and supports the notion of a nonlinear effect of years of education on cognitive performance [34].

A limitation of this study is that a final diagnosis could not be determined in a subset of patients. For example, in some patients it remained unclear whether the symptoms were caused by dementia, a primary psychiatric disorder, other medical conditions (e.g., diabetes and cardiovascular disease), or a combination of these factors. This seems inevitable, as diagnosing dementia in non-Western immigrants is challenging [35, 36] and both under- and overdiagnosis of dementia are common [37]. Strengths of this study were the relatively large sample size and the administration of both the VAT and mVAT within the same person, enabling a direct comparison between the performance on the 2 tests.

In conclusion, the mVAT is a valid, brief, and easy-to-administer test to measure visual association memory in low-educated immigrant populations. Clinicians should be aware that memory performance may be systematically underestimated when using memory tests with line drawings. Future studies with the mVAT should be aimed at including more dementia patients with a wide variety of dementia diagnoses, so the discriminative abilities can be examined for various subtypes of dementia and dementia stages.

Statement of Ethics

The study protocol has been approved by the research institute’s committee on human research.
Disclosure Statement

The authors have no conflict of interests to declare.

Funding Sources

This study was supported by grant 733050834 from The Netherlands Organization of Scientific Research (ZonMw Memorabel).

Author Contributions


References